Exam SØK3001 V24.

Tables for the t-distribution and the F-distribution are included at the end of the sheet.

Permitted examination support material C:
Mathematical manual:
Knut Sydsæter, Arne Strøm og Peter Berck (2006): Matematisk formelsamling for økonomer, 4utg. Gyldendal akademiske.
Knut Sydsæter, Arne Strøm, og Peter Berck(2005): Economists' mathematical manual, Berlin.

Calculator:
- Casio FX-82CW, Casio FC100 V2, Casio fx-82ES PLUS og Casio fx-82EX
- Citizen SR-270X og Citizen SR-270X College

Hewlett Packard HP30S

Question 1 (20%).

Briefly explain the following concepts

a) Serially correlated error term
b) Heteroscedastic error term
c) Stationary time series
d) Measurement error
e) Structural form equation

Question 2 (40%).

A researcher has estimated the demand function for airline seats in the US based on data for 1149 different routes from the year 1997. A simple demand function for airline seats is

(1) $\log(passen) = \beta_0 + \beta_1 \log(fare) + \beta_2 \log(dist) + \beta_3 [\log(dist)]^2 + u_1$

where

$passen$= average number of passengers per day

$fare$ =average airfare (price) in US$

$dist$=route distance in miles

a)If (1) is truly a demand function, what should be the sign of β_1?

Table 1 shows results from different estimated equations based on the data set for these variables from the year 1997. lpassen is $log(passen)$, lfare is $log(fare)$, ldist is $log(dist)$, ldistsq is $[\log(dist)]^2$. *vhat* is the residual from the regression in column (3). The variable $concen$ is a measure of market concentration on the route defined as the share of flights conducted by the largest carrier (company). Table 2 presents descriptive statistics for passen, fare, dist and concen.

b) What is the estimated price elasticity of demand according to the results in Table 1?

c) What is the interpretation of the regression equations in column (2) and (3)?

d) Explain how the equation in column (2) is estimated.

e) Briefly explain the economic argument for using the variable *concen* as an instrumental variable

f) Explain what is meant by weak instruments. Use the results in Table 1 to decide whether you have a weak instrument problem in this case.

g) Explain how you can test whether the price variable, lfare, is exogeneous. Use the results in Table 1 to test the hypothesis that lfare is exogeneous.

h) A commentator argues that one should test for overidentification restrictions when using the instrumental variable approach. Do you agree with the commentator? Explain your answer.

i) Using the results in column (2), describe how demand for seats depends on route distance

Table 1. Estimation results. Estimated standard errors in parentheses. The text under the column number shows the dependent variable used in the regression.

	(1) lpassen	(2) lpassen	(3) lfare	(4) lpassen
lfare	-0.391	-1.174		-1.174
	(0.067)	(0.388)		(0.367)
ldist	-1.570	-2.176	-0.936	-2.176
	(0.629)	(0.726)	(0.272)	(0.687)
ldistsq	0.116	0.187	0.108	0.187
	(0.048)	(0.061)	(0.021)	(0.058)
concen			0.395	
			(0.063)	
vhat				0.810
				(0.373)
Constant	13.230	18.014	6.190	18.014
	(2.100)	(3.217)	(0.890)	(3.042)
Instruments	-	*concen*	-	-
Observations	1,149	1,149	1,149	1,149
Method	OLS	IV	OLS	OLS
R-squared	0.057		0.408	0.061

Table 2. Descriptive statistics. Means, standard deviations, min and max

	Observations	Mean	Standard deviation	Minimum	Maximum
passen	1149	601.042	763.5326	27	7637
fare	1149	173.752	76.30483	37	460
dist	1149	989.745	612.0313	95	2724
concen	1149	0.61254	0.198131	0.192	1

Question 3 (40%)

Politicians are concerned of lack of teachers. There are teacher shortages in several municipalities. The municipalities have to employ teachers who are not formally qualified to be teachers. The teacher union argues that one reason for teacher shortages is high job pressure related to many students in the classroom. In order to reduce teacher shortages, the union has argued that the municipalities should employ more teachers in order to reduce the job pressure.

You are asked to investigate the claim of the union. You are given access to data for municipalities where

- Short = the percent of teachers not formally qualified
- logTeacher = the logarithm of the number of teachers
- logStudents = the logarithm of the number of students
- Central = an index for the centrality of the municipality. The index has the lowest values for rural areas in the periphery and the highest values for the big cities.

You plan to estimate the following model for the situation in the fall of 2019.

(1) $\text{Short}_i = \beta_0 + \beta_1 \log \text{Teachers}_i + \beta_2 \text{Central}_i + u_i$

where subscript *i* denotes municipality and *u* is the error term.

a) What are the necessary assumptions to obtain unbiased estimators by the Ordinary Least Square (OLS) method?

b) Formulate hypotheses for the coefficients in the model (1).

c) The results for the model are presented in column (1) in Table 3. Explain the statistics R-squared (coefficient of determination) and R-squared adjusted.

d) Table 3 presents the estimated coefficients, with standard errors in parentheses. Interpret the findings in column (1). Are they in accordance with your hypotheses?

e) An adviser suggests that you should include the number of students in the model. More teachers do not imply reduced job pressure if also the number of students increases. It is the number of teachers given the number students that should matter. The adviser suggests that you estimate the model

(2) $\text{Short}_i = \beta_0 + \beta_1 \log \text{Teachers}_i + \beta_2 \text{Central}_i + \beta_3 \log \text{Students}_i + u_i$

The results are reported in column (2) in the table. Comment on the findings.

f) The models in column (4) and (5) in Table 3 have imposed restrictions on equation (2). What are the restrictions? Test separately whether the model in column (4) and the model in column (5) are valid restrictions of the model in column (2).

g) You are informed that you can get data for all years from 2010 to 2019. Can you use this extended data set to improve the credibility of your results? How can you specify the empirical model?

h) You are also informed that some municipalities increased the number of teachers in 2017 because they received increased grants from the national government. Some municipalties experienced a substantial change from this year. Can you use this information to improve the credibility of your results? How will you specify the empirical model?

Table 3. Estimation results. Estimated standard errors in parentheses.

	(1)	(2)	(3)	(4)	(5)
logTeachers	0.37 (0.46)	7.79 (2.07)	0.30 (0.45)	-	-
logStudents	-	-7.49 (3.04)	-	-	-
Central	-0.0173 (0.0036)	-0.0060 (0.0047)	-0.0060 (0.0047)	-0.0040 (0.0035)	-
logTeachers – logStudents	-	-	7.49 (2.04)	7.55 (2.04)	9.45 (1.11)
Constant	14.94 (1.31)	24.85 (2.99)	24.85 (2.99)	24.85 (2.99)	26.61 (2.54)
R-squared	0.1252	0.1533	0.1533	0.1524	0.1498
R-squared adjusted	0.1210	0.1470	0.1470	0.1482	0.1477
Observations	412	412	412	412	412

Statistical tables

TABLE G.2

Critical Values of the *t* Distribution

	Significance Level				
1-Tailed:	**.10**	**.05**	**.025**	**.01**	**.005**
2-Tailed:	**.20**	**.10**	**.05**	**.02**	**.01**
Degrees of Freedom					
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
90	1.291	1.662	1.987	2.368	2.632
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

Examples: The 1% critical value for a one-tailed test with 25 *df* is 2.485. The 5% critical for a two-tailed test with large (> 120) *df* is 1.96.

Source: This table was generated using the Stata® function invt.

TABLE G.3a

10% Critical Values of the *F* Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom: 1	2	3	4	5	6	7	8	9	10
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71
90	2.76	2.36	2.15	2.01	1.91	1.84	1.78	1.74	1.70	1.67
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60

Example: The 10% critical value for numerator $df = 2$ and denominator $df = 40$ is 2.44.
Source: This table was generated using the Stata® function invfprob.

TABLE G.3b

5% Critical Values of the *F* Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom: 1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Example: The 5% critical value for numerator $df = 4$ and large denominator df (∞) is 2.37.
Source: This table was generated using the Stata® function invfprob.